AD _____

GRANT NUMBER: DAMD17-96-1-6088

TITLE: Identification of Two Candidate Tumor Suppressor Genes on Chromosome
17p13.3: Assessment of their Roles in Breast and Ovarian Carcinogenesis

PRINCIPAL INVESTIGATOR: Andrew K. Godwin, Ph.D.

CONTRACTING ORGANIZATION: Fox Chase Cancer Center
Philadelphia, Pennsylvania 19111

REPORT DATE: July 2000

TYPE OF REPORT: Annual

PREPARED FOR:

U.S. Army Medical Research and Material Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and
should not be construed as an official Department of the Army position, policy or decision
unless so designated by other documentation.

DTIC QUALITY INSPECTED 4

1

20001116 027

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE July 2000	3. REPORT TYPE AND DATES COVERED Annual (1 Jul 99 – 30 Jun 00)	

4. TITLE AND SUBTITLE Identification of Two Candidate Tumor Suppressor Genes on Chromosome 17p13.3: Assessment of their Roles in Breast...	5. FUNDING NUMBERS DAMD17-96-1-6088

6. AUTHOR(S)

Andrew K. Godwin , Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Fox Chase Cancer Center ~~7701 Burholme Avenue~~ Philadelphia, Pennsylvania 19111	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

This report contains colored photos

12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for Public release; distribution unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

OVCA1, is a candidate tumor suppressor gene, which maps to a region of frequent allelic loss in breast and ovarian cancer at 17p13.3. OVCA1 is mutated in some tumor cell lines, and its protein levels are decreased or lost in nearly 40% of breast and ovarian adenocarcinomas. Expression of low levels of exogenous OVCA1 results in dramatic growth suppression and decreased levels of cyclin D1. OVCA1 codes for a highly conserved protein with no known function but the C-terminal region of OVCA1 interacts with an RNA binding motif protein, named RBM8A. We have recently found that OVCA1 exists in at least two forms: a 48 kDa and a 50-kDa protein. 2-D gel western blotting revealed that OVCA1 is extensively modified post-translationally in tumor cells leading to a dramatic shift in pI. These aberrant modifications result in the p50OVCA1 form and loss of immunoreactive p48OVCA1. Studies of the molecular mechanisms regulating OVCA1 expression and localization, and its interaction with RBM8A are continuing. Overall, our studies indicate that altered expression and/or post-translational modifications of OVCA1 is associated with the development of breast and ovarian tumors and suggest a potentially new mechanism for the inactivation of tumor suppressors in cancer.

14. SUBJECT TERMS Breast Cancer			15. NUMBER OF PAGES 31
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

(   )   Where copyrighted material is quoted, permission has been obtained to use such material.

(   )   Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

(✓)   Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

(   )   In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

(✓)   For the protection of human subjects, the investigator(s) have adhered to policies of applicable Federal Law 32 CFR 219 and 45 CFR 46.

(✓)   In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Insitutes of Health.

_____   _____8/18/00_____
Principal Investigator's Signature   Date

3

Andrew K. Godwin, Ph.D.

Table of Contents

INTRODUCTION:

Breast cancer is the second most common form of cancer in women, striking 1 out of 8 women in their lifetime. Ovarian cancer strikes fewer women but is generally at an advanced stage at the time of detection. Both diseases are controlled by multiple genetic defects, suggesting the involvement of many different genes, including tumor suppressors. According to the two-hit model of Knudson, both alleles encoding for a tumor suppressor must be lost or inactivated in order for cancer to develop. Based on this model, loss of heterozygosity (LOH) of alleles from tumor tissue has been used to suggest the presence of potential tumor suppressor genes.

The short arm of chromosome 17 is one of the most frequently altered regions in human breast and ovarian cancer. One locus of high allelic loss is at 17p13.1, and contains the tumor suppressor gene, *TP53*. However, we and others have shown a second region of LOH distal to the *TP53* gene, at 17p13.3, in breast tumors and ovarian tumors. Genomic abnormalities involving 17p13.3 has also been reported in primitive neuroectodermal tumors, carcinoma of the cervix uteri, medulloblastoma, osteosarcoma, astrocytoma (22), and acute myeloid leukemia and myleodysplastic syndromes, suggesting that a gene(s) on 17p13.3 may play a role in the development of a wide variety of neoplasms, including breast and ovarian cancer.

We have previously defined a minimum region of allelic loss (MRAL) on chromosome 17p13.3 in genomic DNA from ovarian tumors and breast tumors. Positional cloning and sequencing techniques revealed two genes in the MRAL, referred to as *OVCA1* and *OVCA2*, which overlap one another in the MRAL, and have one exon in common (**Figure 1**). Since translation of *OVCA1* does not proceed into the shared exon, the genes encode for completely distinct proteins. The function of OVCA1 and OVCA2 are unknown and their potential role in breast and ovarian oncogenesis has been a major focus of our studies.

BODY:

Progress Report

Year 4

Months 37-48. Specific Aim 3; Complete cloning and sequencing of clones that interact with OVCA1, initiate characterization of function of these new proteins (if not already known); Evaluate effect of OVCA2 expression alone or in combination with OVCA1 on morphology and cell growth, initiate

5

mutational studies of the OVCA1 homologue in yeast, establish transgenic mice with regulatable OVCA1 expression vectors, initiate gene "knockout" studies in immortalized HOSE cells.

OVCA1 protein interactors

OVCA1 codes for a highly conserved protein with no known function. To help determine the function(s) of OVCA1, we used a yeast two-hybrid screening approach to identify OVCA1-associating proteins. One such protein, which we initially referred to as BOV-1 (Binder of OVCA1-1) is 173 or 174 amino acid in length and appears to be a new member of a highly conserved RNA-binding motif (RBM) protein family, that is highly conserved evolutionarily. Northern blot analysis revealed that BOV-1 is ubiquitously expressed and that three distinct mRNA species are expressed, a 1 kb, a 3.2 kb, and a 5.8 kb transcript. The 1-kb transcript is the most abundant and is expressed at high levels in the testis, heart, placenta, spleen, thymus, and lymphocytes. Using fluorescence in situ hybridization (FISH) and the 5.8-kb cDNA probe, we determined that BOV-1 maps to both chromosome 5q13-14 and to chromosome 14q22-23. Further sequence analysis determined that the gene coding the 1 kb and the 3.2 kb transcripts (hereafter referred to as RBM8A) maps to 14q22-23 whereas a second highly related gene coding for the 5.8 kb transcript resides at chromosome 5q13-14 (hereafter referred to as RBM8B). The predicted proteins encoded by RBM8A and RBM8B are identical except that RBM8B is 15 amino acids shorter at its N-terminus. Molecular modeling of the RNA binding domain of RBM8A and RBM8B, based on homology to the sex-lethal protein of Drosophila, identifies conserved residues in the RBM8 protein family that are likely to contact RNA in a protein-RNA complex. The conservation of sequence and structure through such an evolutionary divergent group of organisms suggests an important function for the RBM8 family of proteins. Further details of these studies can be found in the appended publication (Salicioni et al., 2000).

OVCA1 isoforms

We have recently found that OVCA1 exists in at least two forms: a 48 kDa and a 50 kDa protein (**Figures 2 and 3**). Subcellular fractionation studies indicate that $p50^{OVCA1}$ localizes primarily to the cytoplasm whereas $p48^{OVCA1}$ is predominantly found in the organelle fraction with some in the nuclear fraction (**Figure 4**). Immunohistochemical analysis of normal ovaries and ovarian tumors using $p50^{OVCA1}$ or $p48^{OVCA1}$ specific antibodies demonstrated strong nuclear and cytoplasmic staining in the epithelial cells of the normal ovaries and borderline tumors for both antibodies. In contrast, $p50^{OVCA1}$ was observed primarily in the nucleus, with little or no cytoplasmic staining in the vast majority of malignant ovarian tumor cells. Interestingly, no nuclear or cytoplasmic staining was detected in these adenocarcinomas using

6

antibodies exclusive for p48^{OVCA1}. Furthermore, there was not apparent correlation with OVCA1 expression and cellular proliferation (as determined by ki67 staining) and BRCA1 levels in these tumors. 2-D gel western blotting revealed that OVCA1 is extensively modified post-translationally in tumor cells leading to a dramatic shift in pI from ~9 to ~5.5 (**Figure 5**). These aberrant modifications result in the p50^{OVCA1} form and loss of immunoreactive p48^{OVCA1}. Studies are underway to evaluate a large panel of breast and ovarian tumors for postranslational modifications of OVCA1 and to determine the means by which OVCA1 is modified. Overall, our studies indicate that altered expression and/or post-translational modifications of OVCA1 may be associated with the development of malignant tumors of the ovary and breast and suggest a potentially new mechanism for the inactivation of tumor suppressors in cancer.

As an initial step in determining the residue(s) that is potentially modified, we have randomly mutated the OVCA1 cDNA by passaging the cDNA expression vector through mutator bacteria (XLI-RED; Stratagene). As our initial screen we searched for a cDNA encoding a mutant protein that would be detected using both N-terminal antibody (TJ132) and C-terminal antibodies (FC22). We have found that p50^{OVCA1} is detected using TJ132 while the p48^{OVCA1} is detected with FC22 (as indicated above). When we modified the protein by adding a HA-tag we are able to detect the p50^{OVCA1} with both antibodies. We have very recently completed our initial screen of ~330 clones for reactivity to both antibodies. We identified on clone that encoded for a protein of potential significance. Studies are underway to determine what residue is mutated and how this mutation leads to reactivity with the FC22 antibody. Furthermore, we are determining whether overexpression of the mutant is capable of suppressing tumor cell growth as compared to wild-type OVCA1.

Generation of Ovca-/+ and Ovca-/- mice (update on progress and change in direction)

As indicated in the previous progress report, we had recently cloned the mouse *Ovca* gene locus. We determined the intron/exon boundaries for the entire murine *Ovca1* gene and determined that the deletion of exon 10 through intron 12 would produce non-functional Ovca1 and Ovca2 proteins. Therefore, we chose to selectively remove exons 10-12 of *Ovca1* and exon 1 of *Ovca2* to give us the best chance of observing a phenotype in the mouse, since other open reading frames exist in this region (see previous progress report). Furthermore, if future studies warrant, each gene could be selectively "knockout" to determine if one or all contribute to tumor development.

We have designed and assembled our targeting construct, which should result in the recombination shown below (**Figure 6**). As can be seen in the diagram, we have fine mapped the region and have already derived probes that can be used to tell us if there is successful integration. A *DraIII* digest of DNA combined with Probe A should yield a ~6.5 kb fragment if there is correct insertion versus the normal ~6.0 kb germline fragment. Note that Probe A is derived from outside the targeting construct so that it will only detect either the recombinant event or the germline fragment but not the vector itself. If there is an ES cell clone that appears to be correct, we can then verify this integration by an *EagI* digest on the DNA combined with Probe B for the Southern blot. This will yield a ~2.8 kb fragment for a homologous recombination event vs. a >16 kb germline band. Through Genome Systems we obtained 161 ES cells selected for G418 resistance. At the time of this report, none of the ES cells showed proper integration of the KO vector as determined by Southern blotting and PCR analyses. A second round of transfections was initiated on August 15th through Incyte and ES cells will be available on or around September 1, 2000 for screening.

KEY RSEARCH ACCOMPLISHMENTS:

-Reported the identification and structural analysis of human RBM8A and RBM8B: two highly conserved RNA-binding motif proteins that interact with OVCA1.

-Determined that OVCA1 is expressed in two forms: a 48 kDa and a 50 kDa.

-Identified that GRP 78 interacts with a C-terminal bait of OVCA1 in a yeast two-hybrid trap system.

-Created a knockout vector for the *Ovca* locus in mice.

-Initiated screening for chimeras.

REPORTABLE ACCOMPLISHMENTS (related to the tasks outlined in the approved SOW):

Schultz, D.C., Vanderveer, L., Berman, D.B., Wong, A.J., and Godwin, A.K. Identification of two candidate tumor suppressor genes on chromosome 17p13.3. Cancer Res, 56:1997-2002, 1996.

Schultz, D.C., A.H. Prowse, A.H., Godwin, A.K. Characterization of OVCA1, a novel tumor suppressor: evidence for a role in the development of human epithelial tumors. American Journal of Human Genetics 61:443, 1997

Schultz, D.C., Balasara, B., Testa, J.R., and Godwin, A.K. Cloning and localization of a human diphthamide biosynthesis-like protein gene, *hDPH2L2*. Genomics, 52:186-191, 1998.

Oleykowski, C.A., Bronson-Myllins, C.R., Godwin, A.K., Yeung, A.T. Mutation detection using a novel plant endonuclease. Nucleic Acid Research, 26:4597-4602, 1998.

Prowse, A.H., Bruening, W., Schultz, D.C., Godwin, A.K.. Characterization of OVCA1, a novel tumor suppressor: evidence for a role in the development of human epithelial tumors. Proceedings of American Association of Cancer Research 39: 4234, 1998

Bruening, W. Prowse, A.H., Schultz, D.C., Holgado-Madruga, M., Wong, A., Godwin, A.K. Expression of OVCA1, a candidate tumor suppressor gene, is reduced in tumors and inhibits growth of ovarian cancer cells. Cancer Research, 59:4973-4983, 1999.

Prowse, A.H., Bruening, W., Godwin, A.K. OVCA1, a novel tumor suppressor, is aberrantly expressed in ovarian carcinomas. American Journal of Human Genetics 65: 1781, 1999.

Salicioni, A.M., Bruening, W., Vanderveer, L., Godwin, A.K. Functional characterization of OVCA1, a putative tumor suppressor. Am. J. Hum. Genet 65:1797, 1999.

Prowse, A.H., Bruening, B., Godwin, A.K. OVCA1, a candidate ovarian cancer gene, is aberrantly expressed in ovarian carcinomas. (The Fifth Annual Postdoctoral Research Conference, Philadelphia, 2000, oral presentation, award for best presentation).

Salicioni, A.M., Xi, M., Vanderveer, L.A., Balsara, B. Testa, J.R., Dunbrack, R.L., and Godwin, A.K. Identification and structural analysis of human RBM8A and RBM8B: two highly conserved RNA-binding motif proteins that interact with OVCA1, a candidate tumor suppressor. Genomics, accepted, 2000.

Prowse, A.H., Bruening, B., Godwin, A.K. Post-translational Modifications of OVCA1 Contribute to Ovarian Carcinogenesis. (Submitted to American Human Genetics Society, 2000).

Bruening, W., Schultz, D.C., Godwin, A.K. OVCA1, a candidate tumor suppressor, is transiently stabilized during the decision to enter G0 phase. Submitted, 2000.

CONCLUSIONS:

In order for future therapies to be developed for the fight against cancer it is important to understand the basic molecular mechanisms that give rise to a specific cancer type. The fundamental mechanisms underlying the genetic basis of cancer are slowly being defined and involve alterations in genes which have been classified into three general categories: (i) protooncogenes are involved in growth promotion and the defects leading to cancer are a gain of function; (ii) tumor suppressor genes are negative regulators of growth and a loss of function gives rise to cancer; and (iii) DNA repair genes are involved in maintaining the fidelity of the genome and altered function can lead to increase rates of mutations in both classes of cancer-causing genes. Cancer is a multistep process that involves alterations in many specific genes. The normal cell has multiple independent mechanisms that regulate its growth and differentiation and several separate events are required to override these control mechanisms. Progress is now being made in isolating these genes and the proteins they encode for, determining the normal cellular functions of the proteins and in investigating the mechanisms of tumorigenesis.

Breast cancer is a very common disease, causing about 10% of deaths in women in the Western World, while ovarian cancer is the number one gynecologic killer in the United States with over 25,000 diagnosed cases and 14,500 deaths in 1999. Molecular genetic analysis of these tumors has revealed many genetic aberrations that may represent important steps in tumor development. To understand the genetic pathways underlying breast and ovarian tumor development, it is necessary to identify the genes affected by these genetic aberrations and establish any correlations between disruption of their function and tumor phenotype.

Chromosome 17 frequently shows loss of heterozygosity (LOH) in both breast and ovarian carcinomas. In addition, re-introduction of chromosome 17 fragments into breast cancer cell lines has been shown to suppress tumorigenicity. Therefore, inactivation of tumor suppressor genes on chromosome 17 appears to be a critical event in the pathogenesis of breast cancer as well as other cancers. Although *TP53* at chromosome 17p13.1 is involved in the pathogenesis of breast and ovarian cancer, our LOH mapping studies in breast and ovarian carcinomas have defined a region distal to *TP53*, at 17p13.3, thought to harbor

a tumor suppressor gene (Godwin *et al.,* 1994, Schultz *et. al.,* 1996). New genes, *OVCA1* and *OVCA2*, have been identified on chromosome 17p13.3, in this critical region of allelic loss (Schultz *et. al.,* 1996). *OVCA1* is composed of 12 coding exons and one non-coding exon, while *OVCA2* is composed of two exons: a unique exon 1, and an exon 2 which comprises part of the 3' untranslated region of *OVCA1*. Thus, the two genes are overlapping, but their protein products are completely distinct.

Much of our focus during the past few years has been on trying to uncover clues about the function of OVCA1. We have found that OVCA1 is highly conserved and exists in at least three forms; a 48, a 50, and an 85-kDa protein. Evidence suggests that the 85-kDa form is encoded by an alternatively spliced form of *OVCA1* and that the 48- and 50-kDa forms are the result of posttranslational modifications. Subcellular fractionation studies indicate that $p50^{OVCA1}$ localizes primarily to the cytoplasm whereas $p48^{OVCA1}$ is predominantly found in the organelle fraction with some in the nuclear fraction. Western blot analysis revealed that $p50^{OVCA1}$ levels are reduced or are absent in >30% of tumors examined when compared to extracts from normal cells and tissues, and that $p85^{OVCA1}$ is rarely detected in tumors. Somatic mutations have been detected, but are rare in *OVCA1*. Furthermore, two germline missense mutations have been found in breast cancer-prone women who have tested negative for a *BRCA1* or a *BRCA2* mutation. Attempts to create breast and ovarian cell lines that stably over-express the p50 form of OVCA1 have generally been unsuccessful. The clones that do express exogenous $p50^{OVCA1}$ do so at very low levels, and have dramatically reduced rates of proliferation, an increased proportion of the cells in the G_1 fraction of the cell cycle, and decreased levels of cyclin D, which may be caused by an accelerated rate of cyclin D degradation (Bruening, W., *et al.*, 1999). Reversion of these cells to a more rapid growth phenotype is accompanied by complete loss of expression of exogenous OVCA1. Screens for proteins that potentially interact with OVCA1 have uncovered several known and some unidentified proteins, including a novel RNA binding protein (BOV-1/RBM8A) (Salicioni, A.M., *et al.*, 2000). The significance of this interaction is being evaluated and it is known that in eukaryotic cells, different classes of RNA are synthesized in the nucleus, including messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA) and small nuclear RNA (snRNA), and must be actively exported to the cytoplasm. After the transcription of tRNA genes, the resulting RNAs undergo numerous changes before a mature translation-competent species is produced. These changes have been found to include terminal processing, intron splicing, editing, deamination and addition on the nucleotide level (see review by (Ibba and Soll, 1999)), many of which are mediated by RNA-binding proteins (RBPs) and by small RNAs as stable RNP complexes. The most notable feature of the RBPs is the presence of one or more copies of the RRM, the most widely found and best-characterized

RNA-binding motif (Birney, et al., 1993; Burd and Dreyfuss, 1994a). We have determined that RBM8A has only one copy of this RRM with one set of the two consensus nucleic acid binding motifs, RNP-1 and RNP-2, which are characteristic of the heterogeneous nuclear ribonucleoprotein (hnRNP) family of proteins.

The RRM is the only RNA-binding motif for which detailed structural information is available. The structures of several RRM proteins have been determined, and all consist of the same basic protein fold. These include the NH_2-terminal RBD of U1 snRNP A (Tang and Rosbash, 1996), the U1 domain of hnRNP A1 (Xu, et al., 1997), the single RBD of hnRNP C (Gorlach, et al., 1994) , the tandem RRMs of the Sex-lethal protein (SXL) (Crowder, et al., 1999) and the poly(A)-binding protein (PABP) (Deo, et al., 1999). On one hand, SXL governs sexual differentiation and dosage compensation in *Drosophila melanogaster* by binding specific RNA transcripts, while it is known that the mRNA poly(A) tail is an important contributor to both translation initiation and mRNA stabilization/degradation (Sachs and Wahle, 1993). We have built models of RBM8A based on the Sxl and PABP structures. These models identify residues on the surface of the sheet that are likely to bind single-stranded RNA, and share identity in other RNA binding domains with the same protein fold. Since many of these interactions make contact with RNA bases, mutations of these residues are likely to alter the binding affinity of RBM8A for its natural RNA ligand.

In contrast to the RRM, much less is known about the structure and function of the RS and RG motifs or other auxiliary domains found in the RBPs. In considering the possible essential function of RBM8 proteins, sequence analysis of RBM8A allowed us to predict the presence of clusters of arginine/serine-rich and arginine/glycine-rich regions, known as SR and RG motifs, respectively. Furthermore, RBM8A has significant homology to the RG and SR domains described for other human members of the RRM superfamily, i.e. nucleolin, the EWS proto-oncogene and splicing factors, with important roles in mRNA biogenesis and rRNA synthesis, and in the Ewing's sarcoma translocation (Burd and Dreyfuss, 1994b; Delattre, et al., 1992).

An increasing number of proteins appear to be multifunctional, participating in transcriptional and post-transcriptional events. The tumor suppressor WT1, initially considered to be a typical transcription factor (Hastie, 1994), may also be involved in splicing (Lamond, 1995). Interestingly, the N-terminus of WT1 may also contain a cryptic RRM, discovered through molecular modeling (Kennedy, et al., 1996). Furthermore, recent evidence shows that WT1 co-localizes and is physically associated with splice factors,

and suggests that deregulation of splicing may be a crucial factor in the tumorigenesis of the genitourinary system (Davies, *et al.*, 1999). Aberrant splicing has been linked to the early and intermediate stages of mammary tumor formation (Stickeler, *et al.*, 1999), and a deregulated expression balance between hnRNPs and SR factors has been reported for human colon tumor formation and metastasis (Ghigna, *et al.*, 1998).

The importance of our findings is reflected in the recent discoveries that several human and other vertebrate genetic disorders (Buckanovich and Darnell, 1997; Dropcho and King, 1994; Nishiyama, *et al.*, 1998) are caused by aberrant expression of RNA-binding proteins. Furthermore, recent genetic linkage studies suggest that 14q22-23 loci may be implicated in the etiology of human cancers and other diseases (Black, *et al.*, 1999; Knuutila, *et al.*, 1999; Zech, *et al.*, 1999). The conservation of RBM8A sequences through such an evolutionarily divergent group of organisms, from yeast to mice and human, points towards an important function for the RBM8 family of proteins. Our work also provides a starting point for further structural, biochemical and genetic studies of the RBM8 protein family and the biological relevance of its interaction with OVCA1 in the context of normal and abnormal cell growth.

Our studies continue to suggest that OVCA1 has certain properties that are in common with a number of tumor suppressors. We have found that exogenous expression of OVCA1 can inhibit tumor cell growth and that expression of the protein is altered in both breast and ovarian tumors. Yet through our studies, we have not been able to establish a likely function for OVCA1. Therefore, we have initiated studies in mice to evaluate the effect(s) of altering OVCA1 (and in some cases OVCA2) expression on normal growth and development. By establishing such models, we should be better able to identify the function(s) of this very unique, but highly conserved protein.

13

REFERENCES:

Birney, E., Kumar, S. and Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* **21:** 5803-5816.

Black, G. C., Perveen, R., Wiszniewski, W., Dodd, C. L., Donnai, D. and McLeod, D. (1999). A novel hereditary developmental vitreoretinopathy with multiple ocular abnormalities localizing to a 5-cM region of chromosome 5q13-q14. *Ophthalmology* **106:** 2074-2081.

Buckanovich, R. J. and Darnell, R. B. (1997). The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol. Cell. Biol.* **17:** 3194-3201.

Burd, C. G. and Dreyfuss, G. (1994a). Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265:** 615-621.

Burd, C. G. and Dreyfuss, G. (1994b). RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.* **13:** 1197-1204.

Crowder, S. M., Kanaar, R., Rio, D. C. and Alber, T. (1999). Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal. *Proc Natl Acad Sci U S A* **96:** 4892-4897.

Davies, R., Moore, A., Schedl, A., Bratt, E., Miyahawa, K., Ladomery, M., Miles, C., Menke, A., van Heyningen, V. and Hastie, N. (1999). Multiple roles for the Wilms' tumor suppressor, WT1. *Cancer Res.* **59:** 1747s-1750s; discussion 1751s.

Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G. and et al. (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359:** 162-165.

Deo, R. C., Bonanno, J. B., Sonenberg, N. and Burley, S. K. (1999). Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98:** 835-845.

Dropcho, E. and King, P. (1994). Autoantibodies against the Hel-N1 RNA-binding protein among patients with lung carcinoma: an association with type I antid-neuronal nuclear antibodies. *Ann. Neurol.* **36:** 200-205.

Ghigna, C., Moroni, M., Porta, C., Riva, S. and Biamonti, G. (1998). Altered expression of heterogenous nuclear ribonucleoproteins and SR factors in human colon adenocarcinomas. *Cancer Res.* **58:** 5818-5824.

Gorlach, M., Burd, C. G. and Dreyfuss, G. (1994). The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *J. Biol. Chem.* **269:** 23074-23078.

Hastie, N. D. (1994). The genetics of Wilms' tumor--a case of disrupted development. *Annu. Rev. Genet.* **28:** 523-558.

Ibba, M. and Soll, D. (1999). Quality control mechanisms during translation. *Science* **286:** 1893-1897.

Kennedy, D., Ramsdale, T., Mattick, J. and Little, M. (1996). An RNA recognition motif in Wilms' tumour protein (WT1) revealed by structural modelling. *Nat. Genet.* **12:** 329-331.

Knuutila, S., Aalto, Y., Autio, K., Bjorkqvist, A. M., El-Rifai, W., Hemmer, S., Huhta, T., Kettunen, E., Kiuru-Kuhlefelt, S., Larramendy, M. L., Lushnikova, T., Monni, O., Pere, H., Tapper, J., Tarkkanen, M., Varis, A., Wasenius, V. M., Wolf, M. and Zhu, Y. (1999). DNA copy number losses in human neoplasms. *Am. J. Pathol.* **155:** 683-694.

Lamond, A. I. (1995). RNA processing. Wilms' tumour--the splicing connection? *Curr. Biol.* **5:** 862-865.

Nigg, E. (1997). Nucleocytoplasmic transport: signals, mechanisms and regulation. *Nature* **386:** 779-.

Nishiyama, H., Danno, S., Kaneko, Y., Itih, K., Yokoi, H., Fukumoto, M., Okuno, M., Millan, J., Matsuda, T., Yoshida, O. and Fujita, J. (1998). Decreased expression of cold-inducible RNA-binding protein (CIRP) in male germ cells at elevated temperature. *Am. J. Pathol.* **152:** 289-296.

Stickeler, E., Kittrell, F., Medina, D. and Berget, S. M. (1999). Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. *Oncogene* **18:** 3574-3582.

Tang, J. and Rosbash, M. (1996). Characterization of yeast U1 snRNP A protein: identification of th N-terminal RNA binding domain site and evidence that the C-terminal RBD functions in splicing. *RNA* **2:** 1058-1070.

Xu, R., Jokhan, L., Cheng, X., Mayeda, A. and Krainer, A. (1997). Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure* **5:** 559-570.

Zech, J. C., Morle, L., Vincent, P., Alloisio, N., Bozon, M., Gonnet, C., Milazzo, S., Grange, J. D., Trepsat, C., Godet, J. and Plauchu, H. (1999). Wagner vitreoretinal degeneration with genetic linkage refinement on chromosome 5q13-q14. *Graefes Arch. Clin. Exp. Ophthalmol.* **237:** 387-393.

FIGURES

Figure 1. Schematic of the *OVCA* gene locus.

Figure 2. Western blot analysis of OVCA1. Two polyclonal antibodies to OVCA1 have been derived. The N-terminal antibodies (TJ132) was directed against a 13 amino acid peptide (sequence highlighted in blue), while the C-terminal antibodies (FC22) was derived against a fusion protein (sequence highlighted in yellow). Lower panel demonstrates that antibodies are reactive and specific. A2780, ovarian tumor cell line, HIO-117 and -107, human ovarian surface epithelial cell lines.

Figure 3. Western blot analysis of OVCA1 using C- (left panel) and N-terminal (right panel) antibodies. Lower center panel demonstrates that the antibodies detect different molecular weight species of OVCA1 (50 kDa or 48 kDa). A2780, ovarian tumor cell line, HIO-117 and -107, human ovarian surface epithelial cell lines.

Figure 4. Western blot analysis of $p50^{OVCA1}$ and $p48^{OVCA1}$ in subcellular fractions. A2780 cells were separated into P100 and S100 crude fractions. $p50^{OVCA1}$ is expressed primarily in the S100 fraction, while $p48^{OVCA1}$ is expressed predominantly in the P100 fraction. WCL, whole cell lysate.

Figure 5. Two-dimensional western blot analysis of OVCA1 in 293 cells. Protein extracts from 293 cells stably expressing OVCA1 were separated by 2-D gel electrophoresis and OVCA1 was detected using either the C-terminal or N-terminal antibodies. Note the C-terminal antibodies do not detect the majority of forms recognized by the N-terminal antibodies.

Figure 6. Strategy for targeting *Ovca1*. The structure of the normal allele, targeting vector pScrambler/*Ovca*, and targeted allele are shown. Blue boxes labeled neo and DT-A indicate the neomycin resistance gene and diphtheria toxin α-subunit gene cassettes, respectively. Exons are shown as green or red filled boxes, ex1, ex2, etc., for *Ovca1* or *Ovca2*, respectively, whereas other genomic segments are shown in solid lines. Only relevant restriction sites are shown: H=*HindIII*, Dr=*DraIII*, Ea=*EagI*
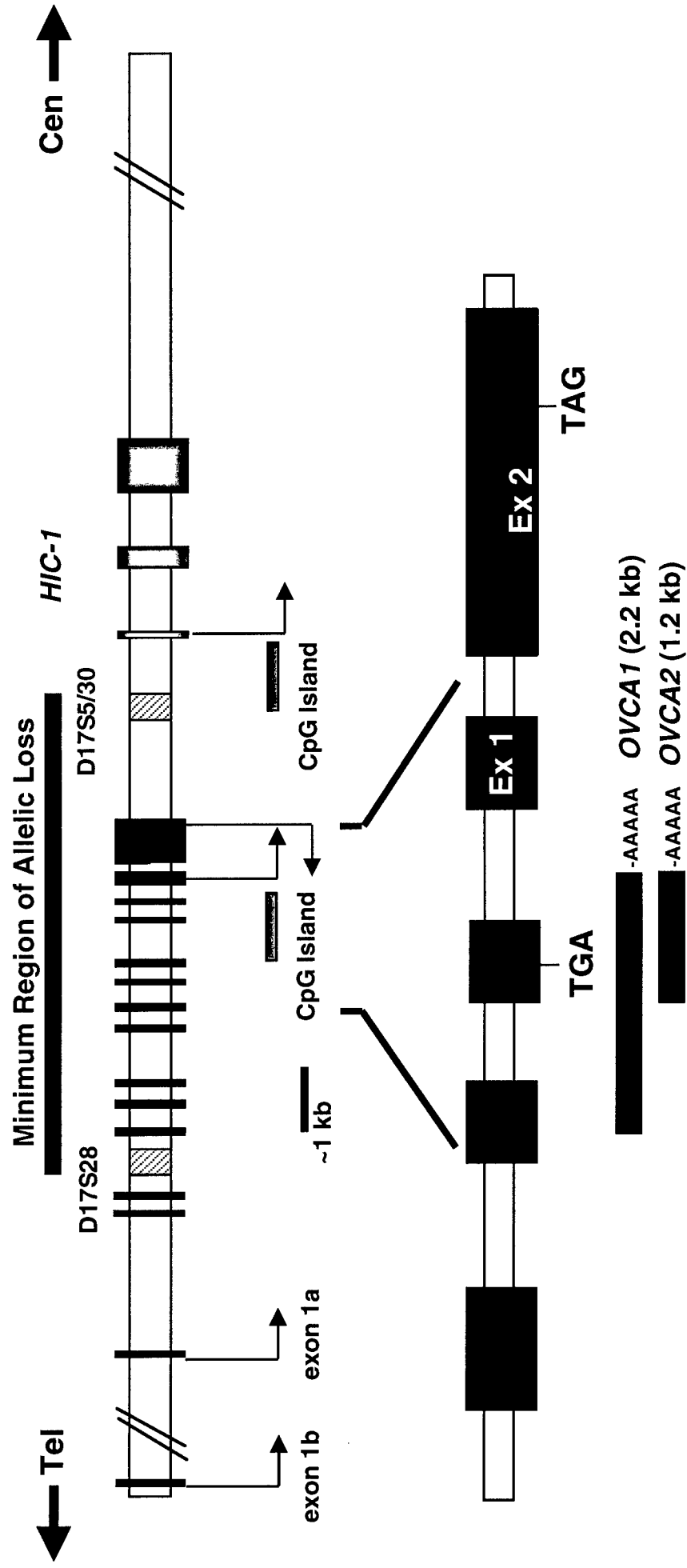
16

Figure 1

Figure 2

MRRQVMAALV VSGAAEQGG   VANQIPPE ILKNPQLQAA IRVLPSNYNF EIPKTIWRIQ QAQAKKVALQ
MPEGLLLFAC TIVDILERFT EAEVMVMGDV TYGACCVDDF TARALGADFL VHYGHSCLMP MDTSAQDFRV LYFVDIRID
TTHLLDSLRL TFPPATALAL VSTIQFVSTL QAAAQELKAE YRVSVPQCKP LSPGEILGCT SPRLSKEVEA VVYLGDGRFH LESVMIANPN
VPAYRYDPYS KVLSREHYDH QRMQAARQEA IATARSAKSW GLILGTLGRQ GSPKILEHLE SRLRALGLSF VRLLLSEIFP SKLSLLPEVD
VWVQVACPRLSID WGTAFPK PLLTPYEAAV ALRDISWQQP YPMDFYAGSS LGPWTVNHGQ DRRPHAPGRP ARGKVQEGSA
RPPSAVACED CSCRDEKVAP LAP

C-terminal antibodies
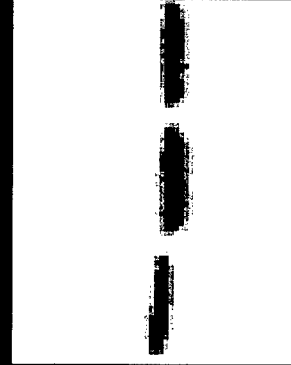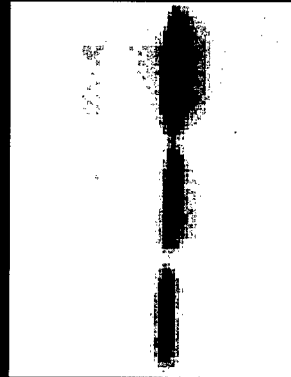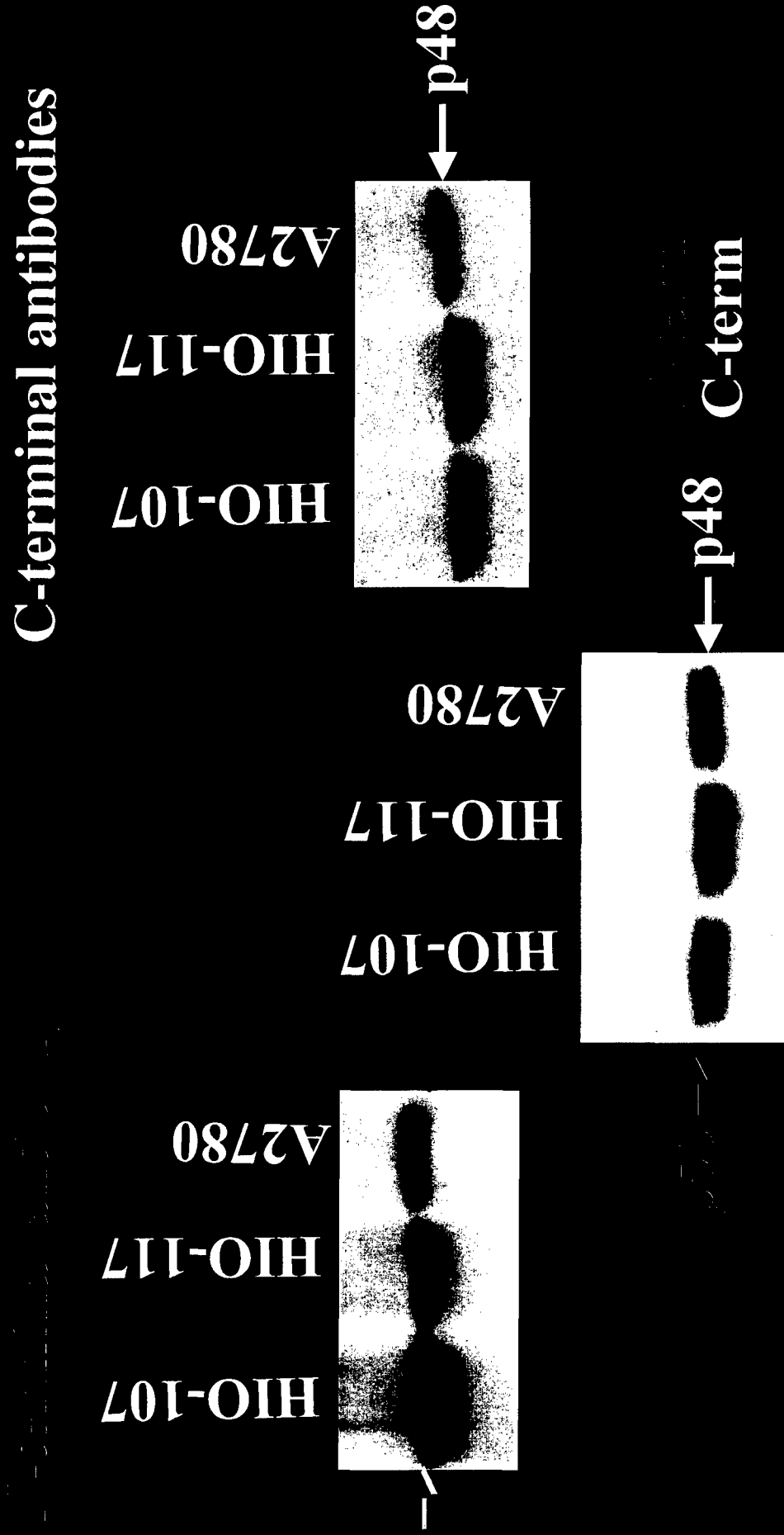
C-term
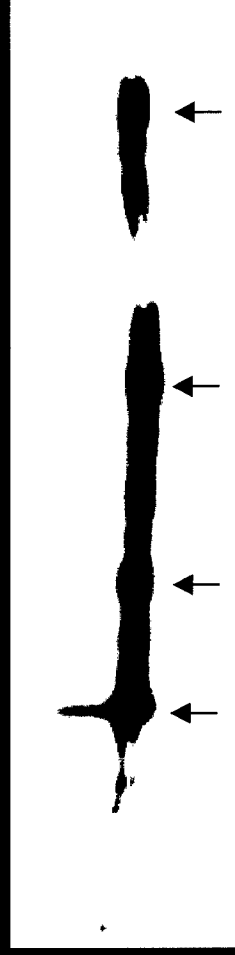
Figure 3

Figure 4

Figure 5
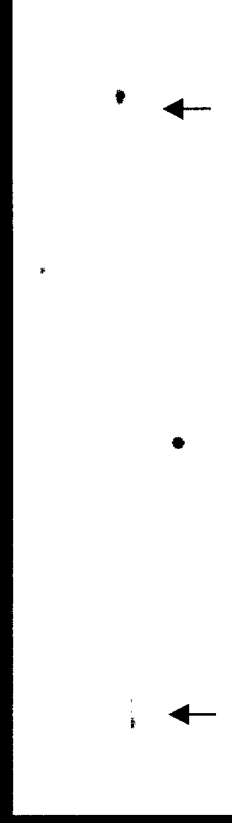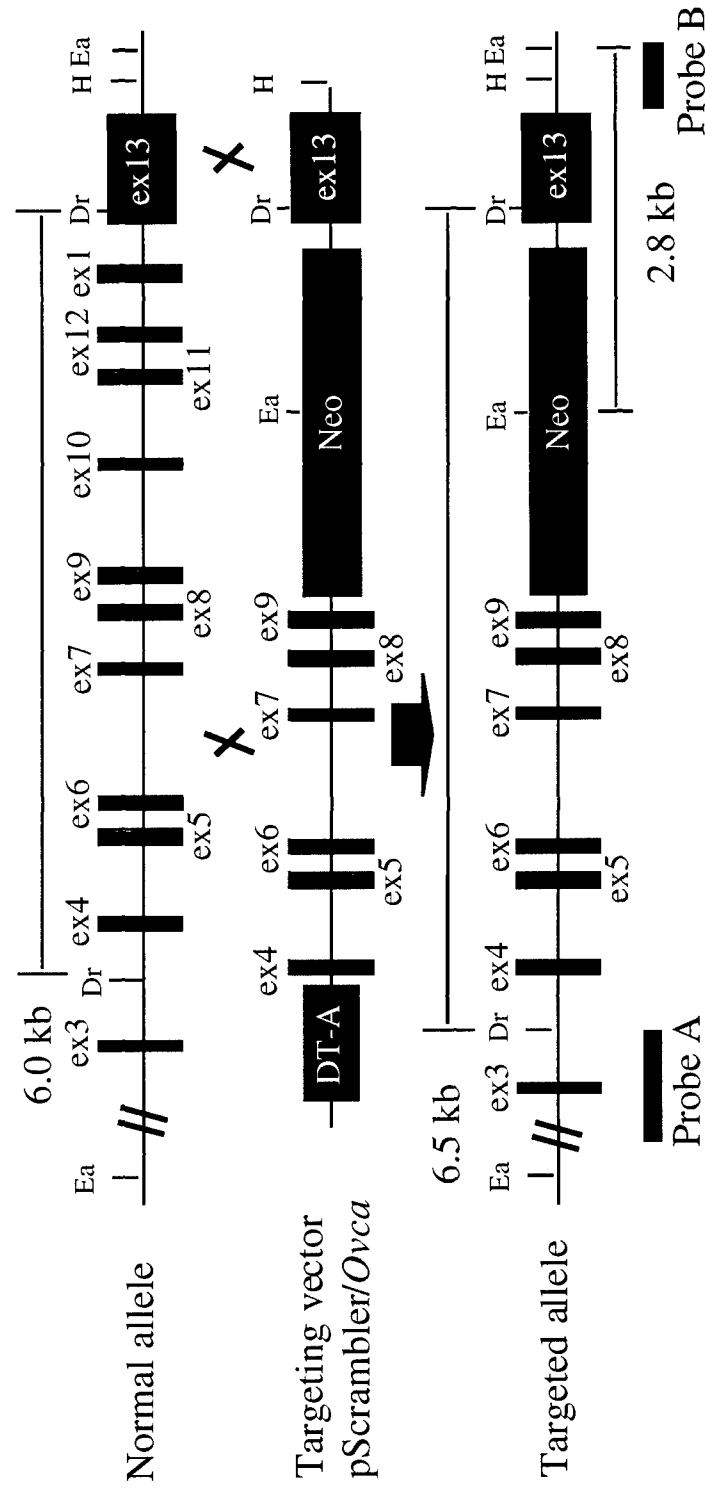
IP 293/OVCA1-HA w/TJ132 → Separated by 2-D gels → Western blotted with N- or C-terminal antibodies

pI 5.4   pI 8.5

C-term

Figure 6

Identification and Structural Analysis of Human RBM8A and RBM8B: Two Highly Conserved RNA-Binding Motif Proteins That Interact with OVCA1, a Candidate Tumor Suppressor

Ana M. Salicioni,* Mingrong Xi,† Lisa A. Vanderveer,* Binaifer Balsara,‡ Joseph R. Testa,‡ Roland L. Dunbrack, Jr.,§ and Andrew K. Godwin*·[1]

*Department of Medical Oncology, ‡Human Genetics Program, and §Basic Science Division, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111; and †West China University of Medical Sciences, Chengdu, Sichuan Province, P.R. China

Received April 6, 2000; accepted June 22, 2000

The *OVCA1* gene is a candidate for the breast and ovarian tumor suppressor gene at chromosome 17p13.3. To help determine the function(s) of OVCA1, we used a yeast two-hybrid screening approach to identify OVCA1-associating proteins. One such protein, which we initially referred to as BOV-1 (binder of OVCA1-1) is 173 or 174 amino acids in length and appears to be a new member of a highly conserved RNA-binding motif (RBM) protein family that is highly conserved evolutionarily. Northern blot analysis revealed that *BOV-1* is ubiquitously expressed and that three distinct messenger RNA species are expressed, 1-, 3.2-, and 5.8-kb transcripts. The 1-kb transcript is the most abundant and is expressed at high levels in the testis, heart, placenta, spleen, thymus, and lymphocytes. Using fluorescence *in situ* hybridization and the 5.8-kb complementary DNA probe, we determined that *BOV-1* maps to both chromosome 5q13–q14 and chromosome 14q22–q23. Further sequence analysis determined that the gene coding the 1- and the 3.2-kb transcripts (HGMW-approved gene symbol *RBM8A*) maps to 14q22–q23, whereas a second highly related gene coding for the 5.8-kb transcript resides at chromosome 5q13–q14 (HGMW-approved gene symbol *RBM8B*). The predicted proteins encoded by *RBM8A* and *RBM8B* are identical except that RBM8B is 15 amino acids shorter at its N-terminus. Molecular modeling of the RNA-binding domain of RBM8A and RBM8B, based on homology to the sex-lethal protein of *Drosophila*, identifies conserved residues in the RBM8 protein family that are likely to contact RNA in a protein–RNA complex. The conservation of sequence and structure through such an evolutionarily divergent group of organisms suggests an important function for the RBM8 family of proteins. © 2000 Academic Press

INTRODUCTION

Analyses of ovarian, breast, and other human tumors indicate that allelic loss of chromosome 17p13.3 is one of the most frequently observed molecular alterations. Recent data from our laboratory suggest that the ovarian cancer 1 gene (*OVCA1*)[2] gene is a strong candidate for the breast and ovarian tumor suppressor gene at chromosome 17p13.3, yet the biochemical function is unknown (Bruening et al., 1999; Schultz et al., 1996). To help determine the function(s) of OVCA1, we used a yeast two-hybrid screening approach to identify OVCA1-associating proteins. Several proteins were identified; however, the most common was a protein containing an RNA-binding motif.

RNA-binding proteins (RBPs) play key roles in the posttranscriptional regulation of gene expression in eukaryotic cells. Once produced in the nucleus, messenger RNAs (mRNAs) are transported to the cytoplasm where the protein synthesis machinery is located, and there is an array of posttranscriptional mechanisms that control mRNA stability, localization, and translation (St Johnston, 1995). Many of these processes are mediated by RBPs and by small RNAs as stable ribonucleoprotein (RNP) complexes. Characterization of these proteins has led to the identification of several conserved RNA-binding motifs (Birney et al., 1993; Burd and Dreyfuss, 1994a). The most notable feature of many of these proteins is the presence of one or more RNA-recognition motifs (RRMs). The RRM is a 90-amino-acid protein domain that binds single-stranded RNA. In contrast to the RRM, much less is known about the structure and function of other RNA-binding motifs, including the arginine/serine-rich (RS) and arginine/glycine-rich (RG) regions. We report here the initial characterization and structural analysis of RBM8A and a closely related member of this highly conserved RNA-binding motif protein family referred

1

to as RBM8B, which interact with a candidate tumor suppressor.

MATERIALS AND METHODS

Yeast two-hybrid interaction trap analysis. The two-hybrid screening of protein interaction was performed by standard protocols (Golemis and Serebriiskii, 1998) using a human fetal brain yeast expression library in the vector pJG4-5 kindly donated by Dr. Erica Golemis (Fox Chase Cancer Center), pMW103 as the LacZ reporter, and an N-terminal or a C-terminal bait, corresponding to amino acids 2 to 161 or 225 to 443 of OVCA1 (i.e., LexA-OVCA1a.a.2–161 or LexA-OVCA1a.a.225–443), respectively.

Cloning of binder of OVCA1-1 (BOV-1). The entire coding regions of *BOV-1a, b,* and *c* were obtained by using an *EcoRI/XhoI* 724-bp complementary DNA (cDNA) fragment, to probe directly a human fetal brain Uni-ZAP XR cDNA library (Stratagene, La Jolla, CA) by standard approaches. DNA from selected phage clones was sequenced.

Northern analysis. The expression of *BOV-1* mRNA in various human tissues was determined by hybridization of an ~800-bp *BOV-1* cDNA fragment to multiple tissue Northern blots, according to protocols previously described (Schultz *et al.,* 1996).

Fluorescence in situ hybridization analysis of BOV-1 cDNA. Fluorescence *in situ* hybridization (FISH) and detection of immunofluorescence were carried out as previously described (Bell *et al.,* 1995). A 5.8-kb cDNA clone corresponding to *BOV-1c* was biotinylated by standard nick-translation methods. Probes were denatured and hybridized to metaphase spreads overnight at 37°C. Hybridized probe was detected with fluorescein-labeled avidin and amplified by addition of anti-avidin antibody (Oncor) and a second layer of fluorescein-labeled avidin. The chromosome preparations were counterstained with DAPI and observed with a Zeiss Axiophot epifluorescence microscope equipped with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) operated by a Macintosh computer workstation. Digitized images of DAPI staining and FITC signals were captured, pseudocolored, and merged using Oncor version 1.6 software.

Mammalian cell culture and transfection procedures. The African green monkey kidney cell line COS-1 (American Type Culture Collection, Manassas, VA) was maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum. The entire coding region of *BOV-1a/b* was PCR-amplified by using oligonucleotides containing both T7 epitope-tag coding sequences and *KpnI/NotI* overhangs and directionally cloned into the *KpnI/NotI* sites of the pCDNA3 expression vector (InVitrogen, Carlsbad, CA), creating a full-length in-frame T7–BOV-1 fusion protein. COS-1 cells were transiently transfected with the indicated plasmid using FuGene (Roche), as directed by the manufacturer, and used for immunofluorescence studies.

Immunofluorescence microscopy. COS-1 cells were seeded onto coverslips in 100-mm dishes and grown to 50% confluence. Forty-eight hours after transfection, cells were fixed in 3% paraformaldehyde and then permeabilized with 2% Triton X in PBS. After incubation with monoclonal antibodies against the T7 tag (Novagen, Madison, WI), cells were washed and then incubated with a FITC-conjugated anti-mouse antibody (Jackson ImmunoResearch Laboratories Inc., West Grove, PA). Slides were visualized with a Nikon Eclipse E800 epifluorescence microscope, and immunofluorescence images were taken with a Nikon camera.

Sequence comparison analysis. Nucleic acid and amino acid homology searches were conducted by using the NetBLAST and FastA programs of the GCG sequence analysis package (Madison, WI) by searching the complete combined GenBank/EMBL and GenBank CDS translations + SwissProt + PDB + SPupdate + PIR databanks, respectively. The expressed sequence tag database at NCBI was searched with the BLAST program.

Homology modeling of the RNA recognition motif domain of BOV-1/RBM8. We used methods described previously to build a model of the RBM8 RRM domain (Dunbrack, 1999). Briefly, PSI-BLAST was used to build a sequence profile of the RBM8 family by iteratively searching the nonredundant protein sequence database available from NCBI (Altschul and Koonin, 1998). We used four iterations of PSI-BLAST, and only sequences with expectation values better than 0.0001 were included in the sequence profile matrix. Upon completion, this matrix was used to search a database of protein sequences in the Protein Data Bank (Berman *et al.,* 2000) of experimentally determined protein structures. This resulted in a list of proteins that could be used as a template for modeling of RBM8A/B. Models of RBM8A/B were built from two of these structures using the side-chain conformation prediction program SCWRL (Bower *et al.,* 1997). SCWRL builds side chains on a template backbone by first placing residues according to a backbone-dependent rotamer library (Dunbrack and Cohen, 1997), followed by a combinatorial search to remove steric overlaps.

RESULTS

Isolation of BOV-1 cDNA

We used a yeast two-hybrid screen to identify cDNAs from a human fetal brain library encoding proteins that were able to interact with OVCA1. A C-terminal bait of OVCA1, corresponding to amino acids 225 to 443, yielded the only protein interactors. A total of 3.5×10^6 primary transformants were screened, resulting in the identification of 28 redundant clones coding for 4 OVCA1 interactor candidates. The most redundant clone, initially referred to as *BOV-1* (binder of OVCA1-1), accounted alone for 54% of the total cDNA isolated. It represented 15 independent isolates of a cDNA of 700 bp in length, encoding a 173-amino-acid protein. The three additional candidate OVCA1 binders were BIP/GRP78 and two previously uncharacterized proteins (data not included).

Northern Analysis of BOV-1

The expression pattern of *BOV-1* mRNA was evaluated by multiple tissue Northern blotting. Three major mRNA species were detected, *BOV-1a, 1b,* and *1c,* of ~1, ~3.2, and ~5.8 kb, respectively. While these species were expressed in all tissues to varying degrees, the 1-kb transcript was most abundant in testis, heart, placenta, spleen, thymus, and lymphocytes (Fig. 1). F1 The three mRNA species, *BOV-1a, 1b,* and *1c,* could also be detected in mammalian cell lines to varying degrees (data not shown).

Cloning of BOV-1 cDNA

To aid in the characterization of *BOV-1,* we isolated 30 cDNA clones from a human fetal brain library using a random-primed 700-bp cDNA probe and sequenced 14 of them to determine the nucleotide sequence and predicted amino acid sequence of *BOV-1a, 1b,* and *1c.* Our results indicate that *BOV-1a* (the abundant ~1-kb transcript) represents the entire coding region identified through the yeast two-hybrid screen and that *BOV-1b* results from the use of an alternative polyadenylation signal (Accession No. AF231511). Based on initial sequence analysis, it appears that *BOV-1c* may
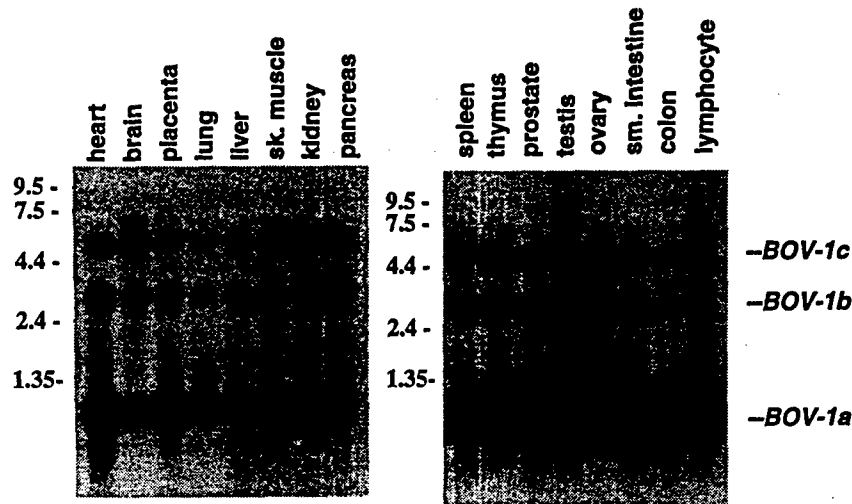
FIG. 1. Tissue expression pattern of BOV-1 mRNA. Blots containing 5 µg of poly(A)$^+$ selected mRNA from each of the indicated human tissues were hybridized with an ~800-bp *BOV-1a/b* cDNA clone. Size standards are in kilobases.

be the product of alternative exon splicing and the use of the alternative polyadenylation signal (Accession No. AF231512).

The cDNA encoding *BOV-1a* consists of 700 bp including 20 bp of 5'-untranslated region (UTR), an AUG leading into an open reading frame (ORF) of 173 amino acids, and a 3'-UTR of 161 bp. The cDNA for *BOV-1b* differs from *BOV-1a* in that the 3'-UTR is substantially longer, 2236 bp versus 161 bp, respectively. The cDNA encoding *BOV-1c* consists of 5786 bp including 3074 bp of 5'-UTR, an ORF of 158 amino acids, and a 3'-UTR of 2238 bp. At the nucleotide level, *BOV-1b* and *BOV-1c*

cDNAs are identical except for the 5'-UTRs (compare sequences for Accession Nos. AF231511 and AF231512).

The predicted protein encoded by *BOV-1c* differs from *BOV-1a/b* in that the protein is predicted to be 15 amino acids shorter (see Fig. 2; translation of protein encoded by *BOV-1c* is predicted to start at the second methionine, but includes an additional amino acid at codon 27); otherwise the proteins are identical. The predicted molecular masses for the BOV-1a/b and BOV-1c proteins are 20 and 18 kDa, and their isoelectric points occur at pH 5.78 and pH 7.62, respectively.
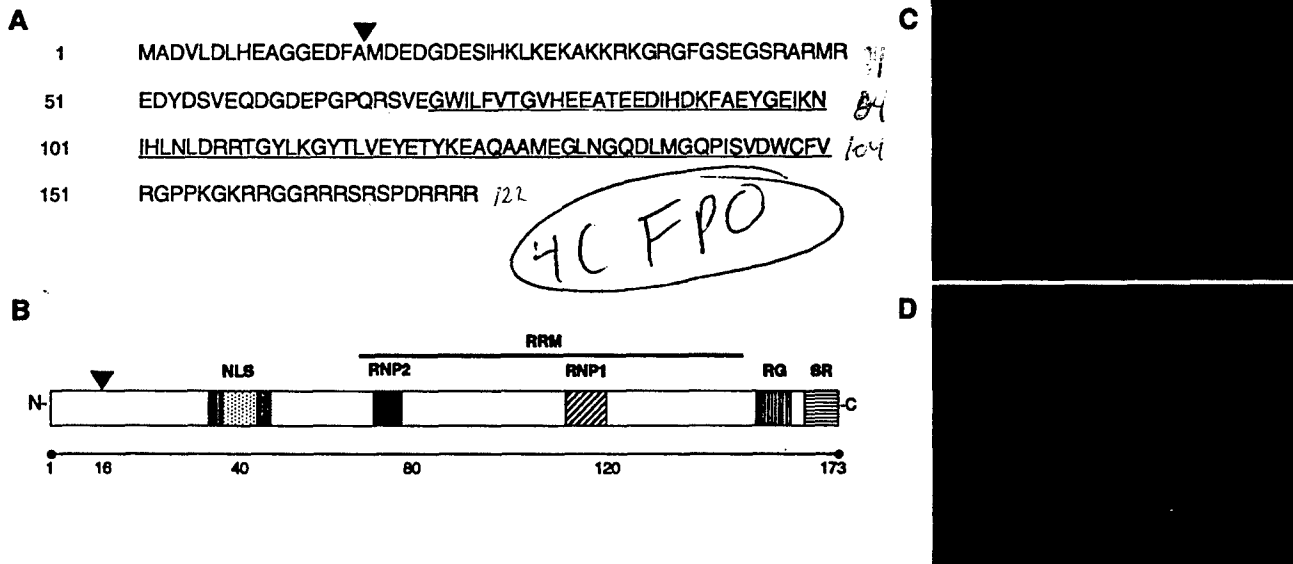

FIG. 2. RBM8A structural (A and B). (A) The predicted primary sequence of the protein encoded by *BOV-1a/b* cDNA is shown, and the RNA-recognition motif (RRM) is underlined. The arrow indicates the methionine where protein encoded by *BOV-1c* is predicted to start. (B) A schematic diagram highlighting structural domains of BOV-1a/b is presented in the lower panel. The RRM and both RNP1 (residues 113–120) and RNP2 (residues 74–79) consensus sequences are indicated. A putative bipartite nuclear localization signal (NLS) is predicted to be present at the N-terminus of BOV-1a/b. The diagram also shows the arginine/glycine-rich (RG) box (residues 151–162) and a serine/arginine-rich (SR) domain (residues 163–173) identified in the C-terminus of BOV-1a/b. The protein encoded by *BOV-1c* is predicted to be 16 amino acids shorter than BOV-1a/b. The arrow indicates the methionine where protein encoded by *RBM8B* is predicted to start. (C and D) Immunolocalization of BOV-1a/b in COS-1 cells. COS-1 cells were transiently transfected with pcDNA3/T7-BOV-1. Forty-eight hours after transfection, cells were fixed and stained with a mouse monoclonal anti-T7 tag antibody (Novagen), followed by staining with FITC-conjugated anti-mouse antibody (Jackson Immunochemicals) (C) and counterstaining with DAPI (Sigma) (D).

Comparison of 12 of the 14 BOV-1 cDNA clones identified a sequence variant involving codon 43 (GAA) of BOV-1a/b (or codon 27 of BOV-1c). In 66.7% (8/12) of the clones, this additional codon was present. These transcripts would be predicted to encode a protein of 174 amino acids (BOV-1a/b). In comparison, this polymorphism was not detected in any of the *BOV-1c* cDNA clones.

Chromosomal Mapping of BOV-1 cDNA

We mapped the chromosomal location of *BOV-1* by FISH using the 5.8-kb cDNA probe, corresponding to the *BOV-1c* transcript. Of the 51 signals observed, 24 (47%) hybridized specifically at 5q13–q14 in 19 of the 20 metaphase spreads scored. In 11 of 20 (55%) metaphase spreads, signals were also detected on chromosome 14, specifically at 14q22–q23. Sixteen (31%) of the 51 signals observed mapped to 14q22–q23, indicating that two closely related genes may exist at these two sites (Fig. 3).

Nucleotide Sequence Analysis of BOV-1 cDNAs

Comparison analysis of *BOV-1a/b* using the BLASTN program demonstrated 99% nucleotide homology (score = 898; *E* value 0.0) to a human EST from a colon carcinoma (HCC) cell line cDNA library (Accession No. AA30779). Comparison of the 5′-UTR of the *BOV-1c* cDNA with the GenBank databases demonstrated 99.4% nucleotide homology over 314 nt (in the reverse orientation) to human integrin-binding protein Del-1 (Del1) mRNA (1712–1399 nt of Del1 and 2668–2981 nt of BOV-1c) (Accession Nos. U70312 and AF231512). The 3′-UTR for both *BOV-1b* and *BOV-1c* also shared nucleotide identity with a human cDNA (Accession No. AL049219) (score = 496; $E = 1.0 \times 10^{-137}$) identified in fetal brain.

Based on comparison of the cDNA sequence for *BOV-*

1a/b and available genomic sequence (Accession No. AF231511), the *BOV-1a/b* gene appears to lack introns, whereas *BOV-1c* contains at least one intron (Accession No. AF231512; and data not shown). The genomic sequencing matching the *BOV-1a/b* cDNA sequence was placed on chromosome 14, further indicating that *BOV-1a/b* (Accession No. AF231511) and *BOV-1c* (Accession No. AF231512) are distinct genes that reside on different chromosomes. Based on our mapping and sequence data, approved names for *BOV-1a/b* and *BOV-1c* have recently been assigned by the Human Gene Nomenclature Committee (http://www.gene.ucl.ac.uk/nomenclature) and are *RBM8A* and *RBM8B*, respectively.

Protein Sequence Analysis and Subcellular Localization of BOV-1/RBM8

The deduced primary amino acid sequence of RBM8A and RBM8B indicates the presence of one copy of an RNA-binding domain (RBD) in the central region (amino acid residues 71–148 or 55–132, respectively), also known as an RRM (Figs. 2A and 2B). This RRM contains one set of the two consensus nucleic acid-binding motifs, RNP-1 (aa 113–120 for RBM8A and aa 97–104 for RBM8B) and RNP-2 (aa 74–79 for RBM8A and aa 58–63 for RBM8B), which are characteristic of the heterogeneous nuclear ribonucleoprotein (hnRNP) family of proteins. The RBM8A and RBM8B amino acid sequence also contains a putative bipartite nuclear localization signal (Robbins *et al.*, 1991) at the N-terminus (aa 33–51 for RBM8A and aa 17–35 for RBM8B) and a stretch rich in glycine residues (not shown). Interestingly, the C-terminus of the RBM8A and RBM8B proteins (residues 151–173 and 135–157, respectively) shows significant homology to the serine/arginine-rich (SR) domain of the splicing factor SC35 (Fu and Maniatis, 1992), as well as a domain rich in glycine and arginine (residues 151–162 for RBM8A and
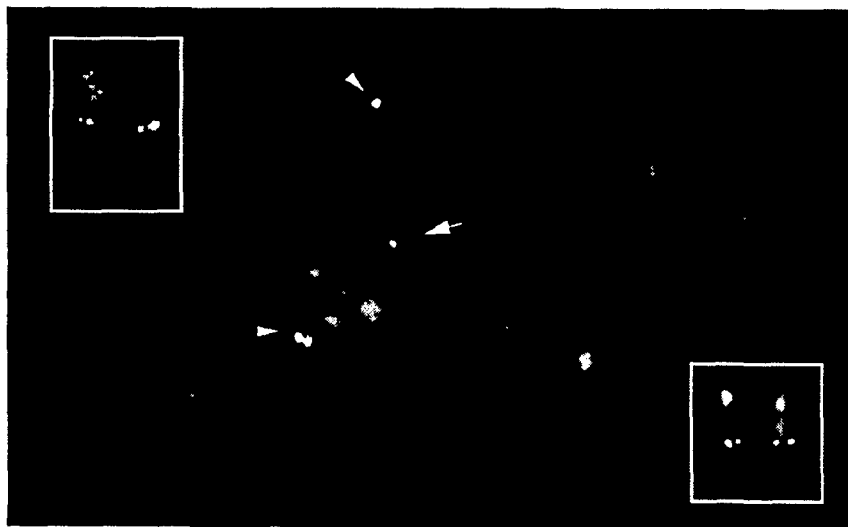
FIG. 3. Chromosomal localization of *BOV-1a/b* and *BOV-1c* by FISH. Partial metaphase spread showing specific hybridization signals at chromosome 5q13–q14 (arrowheads). Arrow indicates specific hybridization signals at chromosome 14q22–q23. (Left inset) *BOV-1c*-specific hybridization at 5q13–q14 to individual chromosomes from other metaphases. (Right inset) *BOV-1c*-specific hybridization at 14q22–q23 to individual chromosomes from other metaphases.

residues 135–146 for RBM8B), reminiscent of the RG box described in human nucleolin (Nigg, 1997). In addition, we analyzed the intracellular distribution of RBM8A by immunofluorescence analysis. Immunolocalization experiments using COS-1 cells transfected with a T7-tagged RBM8A expression vector indicate that full-length RBM8A is localized mostly in the cell nucleus and is finely diffused throughout the cytoplasm (Figs. 2C and 2D).

RBM8 proteins appear to be highly conserved evolutionarily. Sequence analysis of the predicted RBM8A and RBM8B coding sequence using both FastA and TBLASTN algorithms revealed that *Mus musculus,* zebrafish, *Xenopus laevis, Drosophila melanogaster, Caenorhabditis elegans,* and *Schizosaccharomyces pombe* encode hypothetical proteins remarkably similar to RBM8A and RBM8B at the amino acid level (Fig. 4; and data not shown). The percentages of amino acid identity and similarity relative to RBM8A are indicated in the legend to Fig. 4. *M. musculus* hypothetical protein (AL0022712) is 11 amino acids shorter than RBM8A and 100% identical to RBM8A, while zebrafish and *X. laevis,* for example, encode hypothetical proteins with 93.7 and 91.2% identity to RBM8A, respectively. The sequence similarity is spread throughout the entire amino acid sequence and is most pronounced in the RRM region, extending from amino acid residues 70 to 150 of the RBM8A sequence.

RBM8 Protein Model

To gain more information on the RNA-binding properties of RBM8A, we built a three-dimensional model of the RRM domain of RBM8A. Sequence analysis using PSI-BLAST (see Materials and Methods) indicated that there were 18 possible template structures in six families in the Protein Data Bank that could be used to build the RRM domain of RBM8. We chose two of these, the sex-lethal protein from *Drosophila* (Sxl, PDB entry 1b7f) (Handa *et al.,* 1999) and the poly(A)-binding protein (PABP, PDB entry 1cvj) (Deo *et al.,* 1999), since both of these sequences could be aligned with RBM8A without insertions or deletions. The resulting sequence alignment is shown in Fig. 5A. In addition, both structures contained RNA so that interations between RBM8A and RNA could also be modeled.

We built models of RBM8A from both of these structures using the side-chain conformation prediction program SCWRL (Bower *et al.,* 1997). A superposition of the backbones of these two models is shown in Fig. 5B. The root-mean-square (RMS) deviation of the backbones is 0.82 Å. The sequence identity between Sxl and PABP is 31%, while the sequence identity between RBM8 and Sxl is 24% and that between RBM8A and PABP is 27%. Since these sequence identities are similar, we expect that the model of the RRM of RBM8A is quite accurate, with an RMS comparable to the Sxl-PABP RMS of 0.82 Å.

In Fig. 5C, we show the model of RBM8A based on Sxl. Residues that bind to RNA are indicated in Fig. 5C. These include F75, T77, H102, R107, F111, Y115, L117, and P142. Some of these residues are identical or similar in Sxl and/or PABP, which may indicate similar interactions between these side chains and RNA. For example, RBM8A F75 is a tyrosine in PABP and

PRETTYBOX of: 1v1{*} June 8, 2000 11:06:49.78

FIG. 4. RBM8A protein sequence alignment with hypothetical proteins identified in *M. musculus* (100% identity), zebrafish (93.7%/ 98.7%), *X. laevis* (91.2%/95.6%), *D. melanogaster* (63%/76%), *C. elegans* (60.5%/75.3%), and *S. pombe* (48.8%/64.5%), using the UWGCG PileUp program. Percentages of identity/similarity relative to RBM8A are indicated in parentheses. GenBank Accession Nos. AL0022712, AI943400, AW200013, AC006074, CAA83626, and AL021813, respectively. White letters in black boxes indicate identical residues. Shading indicates conserved residues.
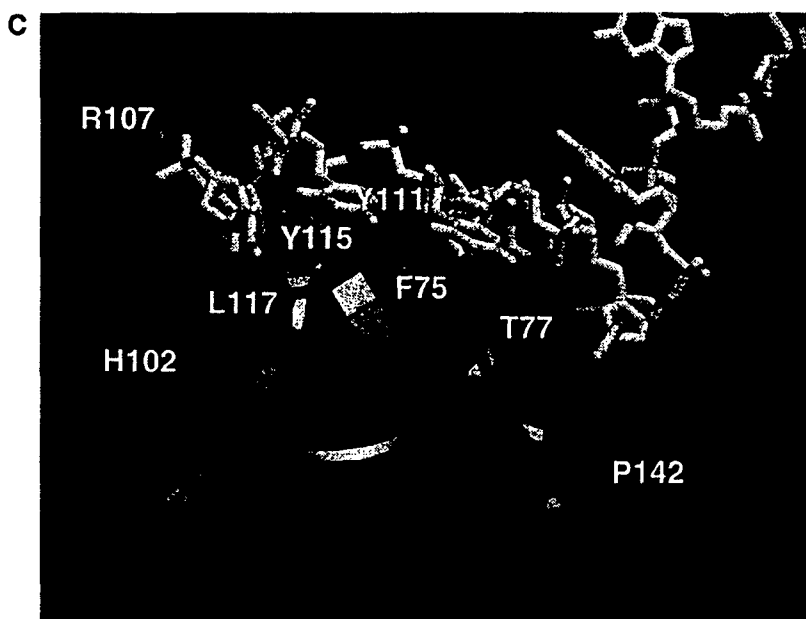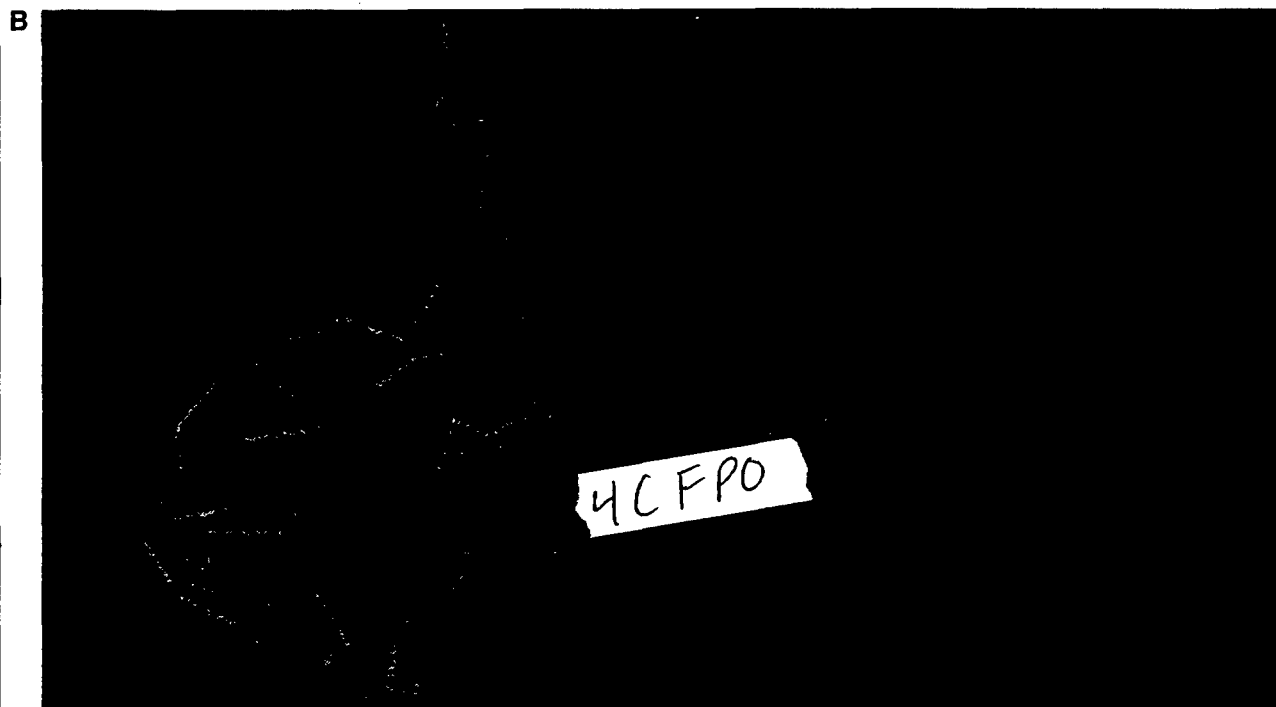
FIG. 5. Homology modeling of the RNA recognition motif domain of RBM8A. (A) Sequence alignment of RBM8A RNA-binding domain with PABP and Sxl from PDB entries 1CVJ and 1B7F, respectively. (B) Superposition of models of BOV-1/RBM8A based on the structures of Sxl (protein in purple and RNA in green) and PABP (protein in yellow and RNA in red). The root-mean-square deviation of the protein backbone Cα coordinates is 0.82 Å. (C) Model of RBM8A based on Sxl in a complex with single-stranded RNA with sequence GUU-GUUUUUUUU. Residues that contact the RNA in either the Sxl or the PABP models are indicated.

forms an aromatic ring-stacking interaction with an adenine in the RNA. In Sxl this residue is an isoleucine, making a hydrophobic interaction with a uracil base. RBM8A Y115 is a tyrosine in both Sxl and PABP, making a ring-stacking interaction with an adenine in

PABP and a hydrogen bond to a uracil ring carbonyl in Sxl. RBM8A Y111 is also a tyrosine in Sxl, which makes a hydrogen bond with a uracil carbonyl in Sxl. While it is not possible to predict the RNA-binding sequence for RBM8A from the model, it is clear that

many of the residues typical of RNA–protein interactions in this family of proteins are contained in the RNA-binding site of RBM8A. Since many of these interactions make contact with RNA bases, mutations of these residues are likely to alter the binding affinity of RBM8A for its natural RNA ligand.

DISCUSSION

Understanding the biological function of the ovarian tumor suppressor OVCA1 may provide new insights into how breast and ovarian cancers develop. To discover important clues into the function of OVCA1, we used a yeast two-hybrid interaction trap method to identify proteins that interact with OVCA1. We identified several candidates (Salicioni *et al.,* unpublished results), including a novel protein that contains a highly conserved RNA-binding motif, initially referred to as BOV-1. In this paper, we demonstrate that *BOV-1a/b* (renamed *RBM8A*) maps to chromosome 14q22–q23 and that the closely related novel gene *BOV-1c* (renamed *RBM8B*) maps to chromosome 5q13–q14. Furthermore, the protein encoded by *RBM8B* is identical to RBM8A except that it is 16 amino acids shorter at the N-terminus and thus appears to be a new member of a highly conserved RBM protein family. Since both RNA transcripts are ubiquitously expressed in human tissues and are predicted to encode a protein, it is unlikely that either gene that we identified is a pseudogene. Two sequences related to RBM8A, (Accession Nos. AF127761 and AF161463) were released in the GenBank database while this article was in preparation. Our results are further supported by the knowledge that genes coding for RNA-binding proteins have diversified by duplication of genes and intragenic domains (Bandziulis *et al.,* 1989; Birney *et al.,* 1993). A three-dimensional model of the RRM domain of RBM8A/B indicates that these RBM8 sequences will fold properly into an RNA-binding domain, forming a hydrophobic core between a β-sheet and two α-helices.

In eukaryotic cells, different classes of RNA are synthesized in the nucleus, including mRNA, ribosomal RNA, transfer RNA (tRNA), and small nuclear RNA, and must be actively exported to the cytoplasm. After the transcription of tRNA genes, the resulting RNAs undergo numerous changes before a mature translation-competent species is produced. These changes have been found to include terminal processing, intron splicing, editing, deamination, and addition on the nucleotide level (see review by Ibba and Soll (1999)), many of which are mediated by RBPs and by small RNAs as stable RNP complexes. The most notable feature of the RBPs is the presence of one or more copies of the RRM, the most widely found and best-characterized RNA-binding motif (Birney *et al.,* 1993; Burd and Dreyfuss, 1994a). We have determined that both RBM8A and RBM8B have only one copy of this RRM with one set of the two consensus nucleic acid-binding motifs, RNP-1 and RNP-2, which are characteristic of the hnRNP family of proteins.

The RRM is the only RNA-binding motif for which detailed structural information is available. The structures of several RRM proteins have been determined, and all consist of the same basic protein fold. These include the NH_2-terminal RBD of U1 snRNP A (Tang and Rosbash, 1996), the U1 domain of hnRNP A1 (Xu *et al.,* 1997), the single RBD of hnRNP C (Gorlach *et al.,* 1994), the tandem RRMs of the sex-lethal protein (SXL) (Crowder *et al.,* 1999), and the poly(A)-binding protein (PABP) (Deo *et al.,* 1999). On one hand, SXL governs sexual differentiation and dosage compensation in *D. melanogaster* by binding specific RNA transcripts, while it is known that the mRNA poly(A) tail is an important contributor to both translation initiation and mRNA stabilization/degradation (Sachs and Wahle, 1993). We have built models of RBM8A based on the Sxl and PABP structures. These models identify residues on the surface of the sheet that are likely to bind single-stranded RNA and share identity in other RNA-binding domains with the same protein fold. Since many of these interactions make contact with RNA bases, mutations of these residues are likely to alter the binding affinity of RBM8A and RBM8B for its natural RNA ligand.

In contrast to the RRM, much less is known about the structure and function of the RS and RG motifs or other auxiliary domains found in the RBPs. In considering the possible essential function of RBM8 proteins, sequence analysis of RBM8A and RBM8B allowed us to predict the presence of clusters of arginine/serine-rich and arginine/glycine-rich regions. Furthermore, RBM8A and RBM8B have significant homology to the RG and SR domains described for other human members of the RRM superfamily, i.e., nucleolin, the EWS proto-oncogene, and splicing factors, with important roles in mRNA biogenesis and rRNA synthesis and in the Ewing sarcoma translocation (Burd and Dreyfuss, 1994b; Delattre *et al.,* 1992).

An increasing number of proteins appear to be multifunctional, participating in transcriptional and post-transcriptional events. The tumor suppressor WT1, initially considered to be a typical transcription factor (Hastie, 1994), may also be involved in splicing (Lamond, 1995). Interestingly, the N-terminus of WT1 may also contain a cryptic RRM, discovered through molecular modeling (Kennedy *et al.,* 1996). Furthermore, recent evidence shows that WT1 colocalizes and is physically associated with splice factors and suggests that deregulation of splicing may be a crucial factor in the tumorigenesis of the genitourinary system (Davies *et al.,* 1999). Aberrant splicing has been linked to the early and intermediate stages of mammary tumor formation (Stickeler *et al.,* 1999), and a deregulated expression balance between hnRNPs and SR factors has been reported for human colon tumor formation and metastasis (Ghigna *et al.,* 1998).

The importance of our findings is reflected in the recent discoveries that several human and other vertebrate genetic disorders (Buckanovich and Darnell, 1997; Dropcho and King, 1994; Nishiyama *et al.,* 1998)

are caused by aberrant expression of RNA-binding proteins. Furthermore, recent genetic linkage studies suggest that both 5q13–q14 and 14q22–q23 loci may be implicated in the etiology of human cancers and other diseases (Black *et al.*, 1999; Knuutila *et al.*, 1999; Zech *et al.*, 1999). The conservation of RBM8A and RBM8B sequences through such an evolutionarily divergent group of organisms, from yeast to mice and human, points toward an important function for the RBM8 family of proteins. Our work also provides a starting point for further structural, biochemical, and genetic studies of the RBM8 protein family and the biological relevance of its interaction with OVCA1 in the context of normal and abnormal cell growth.

ACKNOWLEDGMENTS

REFERENCES

Altschul, S. F., and Koonin, E. V. (1998). Iterated profile searches with PSI-BLAST—A tool for discovery in protein databases. *Trends Biochem. Sci.* **23:** 444–447.

Bandziulis, R. J., Swanson, M. S., and Dreyfuss, G. (1989). RNA-binding proteins as developmental regulators. *Genes Dev.* **3:** 431–437.

Bell, D. W., Taguchi, T., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., Gilks, C. B., Zweidler-McKay, P., Grimes, H. L., Tsichlis, P. N., and Testa, J. R. (1995). Chromosomal localization of a gene, GF1, encoding a novel zinc finger protein reveals a new syntenic region between man and rodents. *Cytogenet. Cell Genet.* **70:** 263–267.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.* **28:** 235–242.

Birney, E., Kumar, S., and Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* **21:** 5803–5816.

Black, G. C., Perveen, R., Wiszniewski, W., Dodd, C. L., Donnai, D., and McLeod, D. (1999). A novel hereditary developmental vitreoretinopathy with multiple ocular abnormalities localizing to a 5-cM region of chromosome 5q13–q14. *Ophthalmology* **106:** 2074–2081.

Bower, M. J., Cohen, F. E., and Dunbrack, R. L., Jr. (1997). Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: A new homology modeling tool. *J. Mol. Biol.* **267:** 1268–1282.

Bruening, W., Prowse, A. H., Schultz, D. C., Holgado-Madruga, M., Wong, A., and Godwin, A. K. (1999). Expression of OVCA1, a candidate tumor suppressor, is reduced in tumors and inhibits growth of ovarian cancer cells. *Cancer Res.* **59:** 4973–4983.

Buckanovich, R. J., and Darnell, R. B. (1997). The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol. Cell. Biol.* **17:** 3194–3201.

Burd, C. G., and Dreyfuss, G. (1994a). Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265:** 615–621.

Burd, C. G., and Dreyfuss, G. (1994b). RNA binding specificity of hnRNP A1: Significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.* **13:** 1197–1204.

Crowder, S. M., Kanaar, R., Rio, D. C., and Alber, T. (1999). Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal. *Proc. Natl. Acad. Sci. USA* **96:** 4892–4897.

Davies, R., Moore, A., Schedl, A., Bratt, E., Miyahawa, K., Ladomery, M., Miles, C., Menke, A., van Heyningen, V., and Hastie, N. (1999). Multiple roles for the Wilms' tumor suppressor, WT1. *Cancer Res.* **59:** 1747s–1750s; discussion 1751s.

Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., de Jong, P., Rouleau, G., *et al.* (1992). Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **359:** 162–165.

Deo, R. C., Bonanno, J. B., Sonenberg, N., and Burley, S. K. (1999). Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98:** 835–845.

Dropcho, E., and King, P. (1994). Autoantibodies against the Hel-N1 RNA-binding protein among patients with lung carcinoma: An association with type I anti-neuronal nuclear antibodies. *Ann. Neurol.* **36:** 200–205.

Dunbrack, R. L., Jr. (1999). Comparative modeling of CASP3 targets using PSI-BLAST and SCWRL. *Proteins Suppl.* **3:** 81–87.

Dunbrack, R. L., Jr., and Cohen, F. E. (1997). Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci.* **6:** 1661–1681.

Fu, X. D., and Maniatis, T. (1992). Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science* **256:** 535–538.

Ghigna, C., Moroni, M., Porta, C., Riva, S., and Biamonti, G. (1998). Altered expression of heterogenous nuclear ribonucleoproteins and SR factors in human colon adenocarcinomas. *Cancer Res.* **58:** 5818–5824.

Golemis, E., and Serebriiskii, I. (1998). Two-hybrid system/interaction trap. In "Cells: A Laboratory Manual" (D. Spector, R. Goldman, and L. Leinwand, Eds.), pp. 69.61–69.40, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Gorlach, M., Burd, C. G., and Dreyfuss, G. (1994). The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *J. Biol. Chem.* **269:** 23074–23078.

Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999). Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **398:** 579–585.

Hastie, N. D. (1994). The genetics of Wilms' tumor—A case of disrupted development. *Annu. Rev. Genet.* **28:** 523–558.

Ibba, M., and Soll, D. (1999). Quality control mechanisms during translation. *Science* **286:** 1893–1897.

Kennedy, D., Ramsdale, T., Mattick, J., and Little, M. (1996). An RNA recognition motif in Wilms' tumour protein (WT1) revealed by structural modelling. *Nat. Genet.* **12:** 329–331.

Knuutila, S., Aalto, Y., Autio, K., Bjorkqvist, A. M., El-Rifai, W., Hemmer, S., Huhta, T., Kettunen, E., Kiuru-Kuhlefelt, S., Larramendy, M. L., Lushnikova, T., Monni, O., Pere, H., Tapper, J., Tarkkanen, M., Varis, A., Wasenius, V. M., Wolf, M., and Zhu, Y. (1999). DNA copy number losses in human neoplasms. *Am. J. Pathol.* **155:** 683–694.

Ladomery, M. (1997). Multifunctional proteins suggest connections between transcriptional and post-transcriptional processes. *BioEssays* **19:** 903–909.

Lamond, A. I. (1995). RNA processing. Wilms' tumour—The splicing connection? *Curr. Biol.* **5:** 862–865.

Nigg, E. (1997). Nucleocytoplasmic transport: Signals, mechanisms and regulation. *Nature* **386:** 779.

Nishiyama, H., Danno, S., Kaneko, Y., Itih, K., Yokoi, H., Fukumoto, M., Okuno, M., Millan, J., Matsuda, T., Yoshida, O., and Fujita, J.

(1998). Decreased expression of cold-inducible RNA-binding protein (CIRP) in male germ cells at elevated temperature. *Am. J. Pathol.* **152**: 289–296.

Robbins, J., Dilworth, S. M., Laskey, R. A., and Dingwall, C. (1991). Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: Identification of a class of bipartite nuclear targeting sequence. *Cell* **64**: 615–623.

Schultz, D. C., Vanderveer, L., Berman, D. B., Hamilton, T. C., Wong, A. J., and Godwin, A. K. (1996). Identification of two candidate tumor suppressor genes on chromosome 17p13.3. *Cancer Res.* **56**: 1997–2002.

St Johnston, D. (1995). The intracellular localization of messenger RNAs. *Cell* **81**: 161–170.

Stickeler, E., Kittrell, F., Medina, D., and Berget, S. M. (1999).

Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. *Oncogene* **18**: 3574–3582.

Tang, J., and Rosbash, M. (1996). Characterization of yeast U1 snRNP A protein: Identification of the N-terminal RNA binding domain site and evidence that the C-terminal RBD functions in splicing. *RNA* **2**: 1058–1070.

Xu, R., Jokhan, L., Cheng, X., Mayeda, A., and Krainer, A. (1997). Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure* **5**: 559–570.

Zech, J. C., Morle, L., Vincent, P., Alloisio, N., Bozon, M., Gonnet, C., Milazzo, S., Grange, J. D., Trepsat, C., Godet, J., and Plauchu, H. (1999). Wagner vitreoretinal degeneration with genetic linkage refinement on chromosome 5q13–q14. *Graefes Arch. Clin. Exp. Ophthalmol.* **237**: 387–393.